

自然さ学習による手書き文字合成

高木, 英行
九州大学大学院芸術工学研究院

ドリンスキー, ヤン
九州大学大学院芸術工学府

<https://hdl.handle.net/2324/4482771>

出版情報 : 電子情報通信学会総合大会 情報・システム講演論文集. 2008, pp.25-, 2008-03-18. The Institute of Electronics, Information and Communication Engineers(IEICE)

バージョン :

権利関係 : Copyright(C) IEICE



自然さ学習による手書き文字合成

Synthesizing Handwritten Characters Using Naturalness Learning

ドリンスキー, ヤン⁽¹⁾

高木英行⁽²⁾

Ján Dolinský

Hideyuki Takagi

九州大学 ⁽¹⁾大学院芸術工学府 ⁽²⁾大学院芸術工学研究院

Kyushu University, ⁽¹⁾ Graduate School of Design, ⁽²⁾ Faculty of Design

1. はじめに

人間の振舞、演奏、書画などと、計算モデルで生成するこれら振舞や演奏との「差」が自然さであると仮定し、(1) 計算モデルからこの自然さを学習するモデルを提案し、(2) この自然さのモデルを手書き文字の自然さ学習に応用し、(3) 手書き文字の自然さの学習にリカレントニューラルネットワーク (RNN) を用いて、RNN の自然さ学習性能の向上について実験的に比較考察する。

2. 自然さのモデル

目標システムの振舞 t は、基本システムの振舞 b と自然さ n との加法的な演算 (優加法も劣加法も含むという意味) で表現される ($t = b \oplus n$) ものと考える。前節の例えを用いれば、人間の演奏は楽譜の正確な音長に自然さの揺らぎを加えることで表現できる、という仮説である。

本論文ではこの仮説を加算演算で単純モデル化する。入力 b から自然さ ($t-b$) を RNN 学習し、RNN 出力の自然さを b に単純加算 ($t = b + n$) することで、目標システムの振舞を合成する。そして、人間の手書き文字 (目標システムの振舞) とフォント (基本システムの振舞) の差分 (図 1 の矢印) を自然さとして学習し、学習した自然さをフォントに加算することで手書き文字合成を行う。なお、本モデルの応用範囲は広く、いろいろな振舞合成に応用可能である。

このモデル化の長所は、 $t = b + wn$ と自然さに重み付け ($0 \leq w \leq 1$) することで、自然さ程度を制御した手書き文字合成が可能になる点である (図 2)。この点が商用の手書きフォント作成と異なる。重み 1 の場合は手書き本人の文字に非常に近い手書き文字を合成し、重み 0.6 ならばフォントから 60% 程その手書き本人の文字に近い手書き文字を合成し、重み 0 はフォント文字そのものを生成する。心理実験の結果、重み係数と視覚的な類似度は線形ではなく多少 sigmoid 的關係にあり、心理実験で得られたこの関係を予め内蔵することで、視覚的な類似度を制御できる。

3. 自然さの NN 学習

フォントストローク情報を図 1 のフォント上の点で示した連続する m 点ベクトルで表現して入力とし、図 1 の矢印で示した手書き文字との差分ベクトル m 点を出力するよう学習することを考える。この入力 m 点と出力 m 点の関係を正準相関分析で解析すると、入出力が線形に近い文字もあったが、多くの文字では入力が表現できる出力の自然さはわずか 30% 程度しかなかった。

このことは、入出力関係の学習には非線形学習手法が不可欠であることを示しており、本論文では手書き文字の自然さをモデル化するため、echo-state 付き RNN [1] を用いることにする。

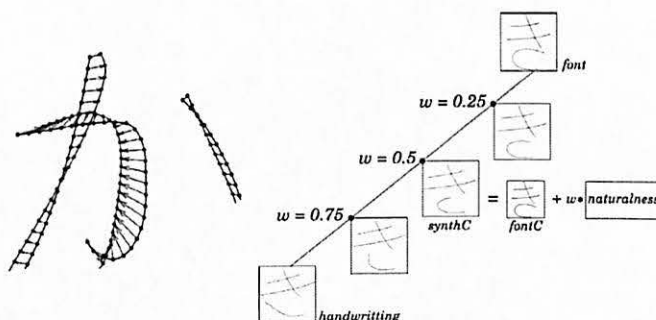


図 1 矢印表現された手書きとフォントの差分ベクトル場

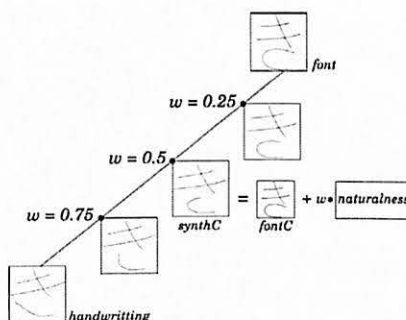


図 2 自然さの重み付け加算によって手書き合成文字の手書き度合いを制御可能

あいうえおかきく こけさしすせそら
りるろなにぬね のはひへほふたち
つてとよゆをわま みむめもばびぶべ
ぼがぎぐげご

図 3 手書き合成文字: ROL なし (左) とあり (右)

4. 実験結果と考察

いろいろな RNN 構造で実験した結果、出力層に自己フィードバックを持つ ROL (recurrent output layer) を導入することによって、自然さ学習性能を向上できることがわかった (図 3) [2]。従来、RNN 性能の低下や不安定化で難色を示されてきた ROL であるが、本実験結果は ROL の安定した実用性を示している。文献[3]に学習の詳細を示す。

謝辞 本研究は一部文科省 21 世紀 COE プログラムの補助を受けた。

参考文献

- [1] H. Jaeger, "The "echo state" approach to analyzing and training recurrent neural networks," Sankt Augustin: GMD-Forschungszentrum Informationstechnik, GMD-Report 148 (Dec., 2001).
- [2] J. Dolinský and H. Takagi, "RNN with a recurrent output layer for learning of naturalness," Int. Conf. on Neural Information Processing, Kitakyushu, Japan (Nov., 2007).
- [3] J. Dolinský and H. Takagi, "Synthesizing Handwritten Characters Using Naturalness Learning," Int. Conf. on Computational Cybernetics (ICCC2007), Gammarth, Tunisia, pp.101-106 (Oct., 2007).